

# A collaborative model for web archiving ephemeral art resources at the New York Art Resources Consortium (NYARC)

Sumitra Duncan

Karl-Rainer Blumenthal

The vast expanse and volatility of art ephemera based on the World Wide Web pose significant threats to the completeness of the art historical record as sustained by art libraries. Towards its mission to enhance the resources available for current and future research through collaboration among leading museum libraries, the New York Art Resources Consortium (NYARC) collects, preserves, and provides access to art ephemera born in digital formats native to the web. It leverages its member institutions' traditional collecting strengths and combined resources to establish an initial and model a permanently sustainable web archiving programme. This article introduces NYARC's web archiving practices as they manifest at the principal stages in a typical web archive's lifecycle, describes how each directly benefits from collaboration among its member libraries and external programme partners, and identifies opportunities for further art libraries and their consortia to participate in this important effort to serve and preserve at-risk art historical resources.

## Introduction: Collecting ephemeral art resources at NYARC

The research libraries and archives of the Brooklyn Museum, The Frick Collection, and The Museum of Modern Art, which comprise the New York Art Resources Consortium (NYARC),

have long demonstrated a commitment to collecting, preserving and providing access to ephemeral materials pertinent to art historical scholarship. The Artist Files collection at The Museum of Modern Art Library alone contains upwards of 80,000 files, inclusive of exhibition announcements, press releases, clippings, brochures, small exhibition catalogs, checklists, invitations, and various other ephemera.<sup>1</sup> The NYARC libraries are widely renowned for offering a wealth of resources, inclusive of art ephemera, of which a great percentage are entirely unique among similar institutions.

Art librarians tasked with collection development have witnessed art ephemera drift to a born-digital format, having historically been produced only in a print form (albeit scarcely), and to formats that are frequently only accessible online for a short period of time. In her article, 'Online Art Ephemera: Web Archiving at the National Museum of Women in the Arts,' Heather Slania refers to 'online art ephemera' as 'websites as well as the information contained within a website in any format, including video and audio.'<sup>2</sup> Given the rapid pace at which websites and their embedded pages or files are updated, transformed and redirected, or simply cease to exist, it is critical that those tasked with collecting the rich art ephemera being produced in the digital age understand the overly transient nature of websites and plan their collection development programmes accordingly.

The NYARC libraries were quick to identify the urgent need to address the collection of born-digital art resources and by 2010 the consortium was proactively pursuing web archiving as a potential solution for collecting ephemeral art-rich websites and their embedded files. Web archiving is the practice of harvesting websites, as they appear on the live web at a given point in time (generally through the use of a web crawler), with storage of the web capture in a standard file format for future preservation and access. The WARC file format, or Web ARChive file, combines multiple digital resources into an aggregate archived file with related metadata.

NYARC's collaborative web archiving programme was effectively established during 2013-15, with the generous support of a two-year grant, 'Making the Black Hole Gray: Implementing the Web Archiving of Specialist Art Resources,' from The Andrew W. Mellon Foundation.<sup>3</sup> NYARC's 2013 grant from the Mellon Foundation supplied the necessary funds to initiate the consortium's subscription to the Archive-It service, hire a full-time NYARC Web Archiving

Coordinator, Sumitra Duncan, to manage the consortium's collaborative web archiving programme, and to employ three part-time paid graduate student interns to conduct quality assurance to NYARC's archived websites, all for a two-year period. Archive-It, the subscription-based web archiving service of the Internet Archive, works with over 400 partner organizations to aid them in harvesting, building, and managing collections of born-digital content from the World Wide Web.<sup>4</sup> The consortium's Archive-It collections and programme-specific documentation are publicly accessible online.<sup>5</sup> While web archiving remains a daunting practice due to the challenges of scale, cost, and ever-evolving web technologies, NYARC has demonstrated that their collaborative approach to web archiving of born-digital art resources allows them to nimbly develop a scalable and sustainable workflow for curation, description and access, and preservation of their web archived content.

#### *National Digital Stewardship Residency: Web Archive Management at NYARC*

As relative newcomers to collecting and preserving born-digital resources at such a large scale, NYARC elected in 2014 to supplement its grant-funded web archiving programme with concurrent and specific guidance on issues of digital preservation. It applied and was selected to host one of five in the first cohort of National Digital Stewardship Residencies (NDSR) in New York.<sup>6</sup> Designed by the Library of Congress's Office of Strategic Initiatives and funded by a grant from the Institute of Museum and Library Services (IMLS), NDSR matches recent graduates of Masters degree programmes in library science and affiliated disciplines with short-term (nine-month) projects of particular significance to the development and dissemination of sustainable digital preservation practices. NYARC's resident, Karl-Rainer Blumenthal, worked with the NYARC web archiving programme's stakeholders across its three institutions and on a full-time basis from September 2014 through May 2015 in order to implement standards and workflows that ensure the viability of NYARC's web archives at their point of collection and for the longer term in their eventual storage and preservation environment, and to do so in such a way that modeled best practices for curators more widely at these critical points in their web archiving processes.

## NYARC's collaborative web archiving lifecycle model

The ongoing work of building and maintaining NYARC's web archives can be defined in terms of a four-stage lifecycle that broadly applies to all such programmes that collect, preserve, and facilitate access to web-native resources:

1. Collection development and curation
2. Harvesting and quality assurance
3. Storage and preservation
4. Description and access

Each stage in this process presents its own opportunities for institutional distinction as well as collaboration. The technical challenges, resource deficiencies, and general curatorial sophistication that may otherwise halt a nascent programme at any of these stages are managed and overcome through collaboration among the NYARC partners and with their broader communities of professional colleagues in both art librarianship and web archiving.

### *Collection development and curation*

A significant component of NYARC's mission is to facilitate collaboration that results in enhanced resources to research communities, with web archiving of particular imperative as additional ephemeral art materials drift to solely born-digital formats. In this case, NYARC has developed a collaborative collection development policy for websites and consortial workflow (see Fig. 1).<sup>7</sup> Websites of scholarly value are selected and nominated for inclusion in the consortium's ten curated Archive-It collections, each of which is aligned with the collecting objectives and strengths of the three NYARC libraries. In order to be pursued for web archiving, nominated websites must align with one of the existing NYARC Archive-It collections, which presently encompass:

1. NYARC institutional websites
  - a. Brooklyn Museum ([brooklynmuseum.org](http://brooklynmuseum.org))
  - b. The Frick Collection ([frick.org](http://frick.org))

- c. Museum of Modern Art (MoMA.org)
  - d. NYARC (nyarc.org)
2. Scholarly art resources deemed at risk of impermanence
  3. Artists' websites
  4. Auction houses
  5. Catalogues raisonnés
  6. New York City galleries and art dealers
  7. Scholarship for the restitution of lost or looted art

Due to the rapid pace of change in the content, functionality, and features of websites, the collection development policy for websites is reviewed and revised accordingly on a periodic basis. Designated staff selectors at each of the NYARC libraries nominate websites for inclusion in the web archives via a simple online form. Additional staff and registered library patrons of the NYARC institutions, as well as the public, may also recommend websites in the fields of their expertise for inclusion. All submissions received from the public or registered library patrons are reviewed by NYARC staff selectors for final approval. Selectors may also make note of content within a nominated website that warrants item-level cataloging, in order to create access to specific PDFs or ephemeral files within a site, such as press releases, exhibition pamphlets, brochures, or catalogs.

Prior to archiving a website, NYARC's web archiving staff contacts the owner of the site to seek their permission to include their content in the consortium's web archive collections. The practice of seeking permission prior to archiving a site was adopted at NYARC after collaborative discussions with the web archiving staff at Columbia University Libraries (CUL), as the more experienced CUL web archiving team had found this to be a sound approach for their own collection development objectives.<sup>8</sup> The structural makeup of a website is additionally evaluated to determine how well the site is likely to capture upon first attempt at archiving. A permissions letter is sent by NYARC, followed by a second letter if no response is received, and should the site owner fail to respond, NYARC sends a third piece of correspondence to notify the site owner of the intent to archive the website (also modeled on CUL's approach). Website owners retain the right to request that their site not be included in the NYARC web archive at any time, whether prior to the initial web crawl or in the future. To date, the vast majority of site owners have enthusiastically provided their permission to NYARC to include their websites in

the web archive collections, as the owners appreciate the great need to preserve the potentially quite ephemeral content that they publish via their sites. For those few site owners that decline permission to be archived, NYARC will generally still provide online access to the live site via the creation of a bibliographic record, as long as the live site continues to exist and function properly.

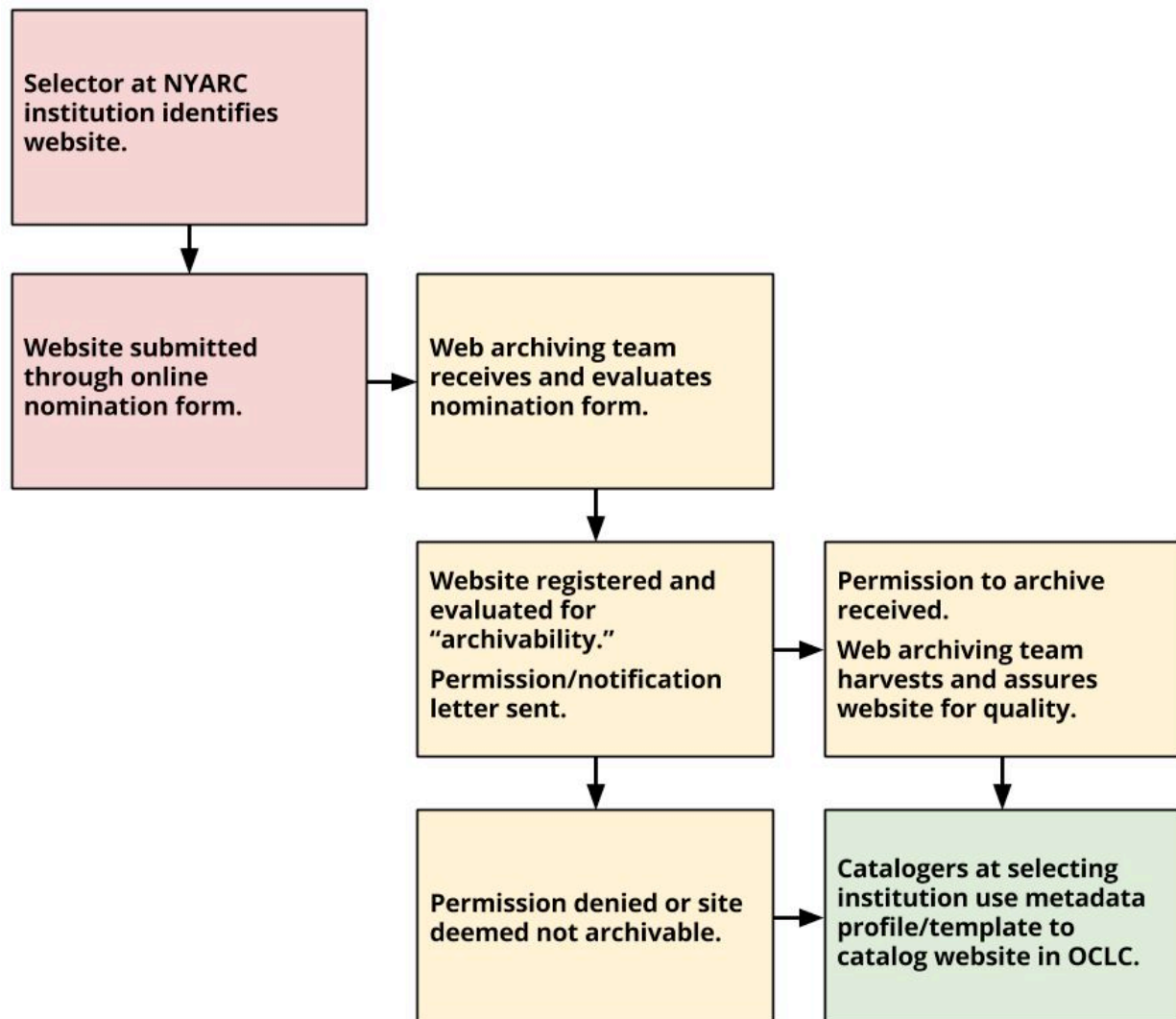


Fig. 1. NYARC's workflow for nomination and cataloging of websites.

### *Harvesting and quality assurance*

Once selected for inclusion in NYARC's collections, websites must be *harvested* (downloaded from the web and into the WARC files over which the collector maintains authority) and assured for their quality such that they may be loaded in a replay mechanism for patron access. As an Archive-It partner, NYARC relies upon long established and widely adopted tools and resources at both ends: the Internet Archive's web crawling technologies in order to harvest websites, and a derivative of and complement to its well known Wayback Machine in order to replay them.<sup>9</sup>

The exponentially increasing dynamism of the web's design and architectural conventions is the predominant challenge to using both of the above tools, and thusly the primary impetus to perform assiduous quality assurance upon newly archived websites. The Internet Archive's web crawling and WARC replay technologies, and the many derivatives that have followed since the introduction of the Wayback Machine in 1996, were initially designed to archive a World Wide Web that was more static and text-based than the one of the present day. While they continuously develop and improve to meet new challenges, these tools nonetheless therefore also struggle to completely harvest and faithfully represent the more user-responsive and visually sophisticated websites<sup>10</sup> of especial interest to stewards of visual culture like art libraries.<sup>11</sup> NYARC routinely confronts these challenges as it archives the sites of artists and their exhibitions, its own institutional online tours and collections databases, and other similarly critical collecting priorities (see Fig. 2). On a select basis, NYARC employs the further services of commercial web archiving vendor Hanzo Archives to harvest these resources.<sup>12</sup> For the vast majority of resources, however, it performs extensive quality assurance in order to guarantee complete and accurate replay.

Quality assurance is the manual process reviewing and, when necessary, improving the completeness and accuracy of harvested websites. It begins with identifying missing content (web pages, images, downloadable documents, etc.) and/or functionality (navigation, responsive scripts and applets, etc.). Archive-It enables NYARC to address such problems of quality with two key resources exclusive to its service: 1) the technical capacity to "patch crawl" and thusly incorporate content detected as missing from web archives when they are first replayed in its version of the Wayback Machine, and 2) the human and technical engineering resources

necessary to address systemic issues that impede its crawling and replay technologies from fully capturing and/or rendering websites' proper behaviors. Both significantly enhance NYARC's ability to provide faithful archival representations of once live websites, but are challenging to maintain at even the relatively selective scale at which it collects. The selection, judgment, and diligence required to improve the quality of websites within NYARC's collecting scope can be difficult to predict. Subsequently, it can be equally difficult to plan and efficiently manage the time and personnel that the process requires.

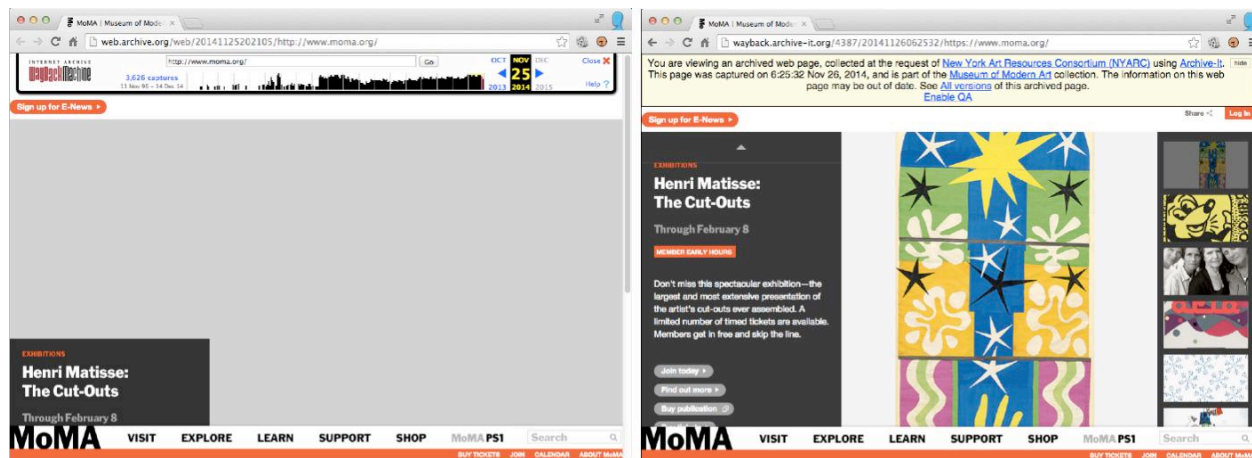


Fig. 2. Two archival renditions of the Museum of Modern Art's homepage--MoMA.org. *Left:* As archived by the Internet Archive's fully automated web crawler on November 25, 2014, and replayed through its Wayback Machine. *Right:* As archived on November 26, 2014, assured for quality, and replayed by NYARC using Archive-It.

To address the first issue of scale, NYARC delegates quality assurance responsibilities among its programme's graduate student interns. Each reviews and improves the archived website of his or her host institution and of a selection of archived websites from among NYARC's seven other thematic collections. This delegation of responsibilities can, however, itself be challenging to manage efficiently. As interns perform their duties for a limited time and dispersed across three host sites, opportunities to consolidate complementary work and retain institutional knowledge of issues and best practices can be lost.

In response to this challenge, NYARC tasked its National Digital Stewardship Resident to survey the quality assurance work of its programme's interns as well as the comparable practices of its colleagues in the web archiving field. By interviewing and benchmarking the quality assurance

practices of peers in art libraries, national libraries, academic special collections, and other web archiving organizations, NYARC was able to define the boundaries and best workflows for the process moving forward. The result, first published to the web in December 2014, is a central point of reference and reporting for NYARC's ongoing quality assurance process.<sup>13</sup> It includes a step-by-step guide through the manual process, the programme's accumulated knowledge of quality issues and their respective improvement strategies, and an online reporting system that empowers all programme staff responsible for quality assurance to summarize their observations, actions, and recommendations for each archived website according to actionable, form-based options.<sup>14</sup> Opportunities to improve the capture and/or replay of website in NYARC's collections may then be prioritized and planned according to available resources.

### *Storage and preservation*

As soon as they are crawled on the live web, all of the constituent components of each selected website are deposited into digital files conforming to the internally standardized and supported Web ARChive (WARC) format.<sup>15</sup> These WARC files are stored on the Internet Archive's online access servers and replayed through Archive-It's and/or Internet Archive's "general" Wayback interface whenever a patron requests access to an archived website through their web browser.

Files in this or any such digital repository can be negatively affected by viruses, data corruption, and other human/and or hardware management failures, thus potentially negating all of the above work to preserve these already exceptionally ephemeral resources. The Internet Archives protects NYARC's and all of its Archive-It partners' WARC files against these threats in three critical ways:

- **Redundancy and geographic distribution:** Both a primary and a 'mirrored' (duplicate) copy of each WARC file is stored in the Internet Archive's data repository in Redwood City, California, 35 miles south of its San Francisco headquarters. For further redundancy and protection against geographically specific threats to this storage location, a third copy of each is also stored in its secondary data repository in San Francisco, and a fourth is periodically shipped on disk to an offline (or 'dark') archive at Old Dominion University

in Norfolk, Virginia. Finally, authorized users of Archive-It software may download and store some or all of their organization's WARC files for on-site preservation by using either a simple web or more advanced command line interface.<sup>16</sup>

- **Data integrity:** Upon introduction into its storage environment, each WARC file is immediately assigned a cryptographic hashing algorithm (or 'checksum') value, with which the Internet Archive's automated scanning systems may monitor its *fixity* ('the property of a digital file or object being fixed or unchanged').<sup>17</sup> The WARC effectively produces a unique fingerprint based upon its precise contents, and, should that fingerprint ever change due to data loss or corruption, the system can alert the WARC's owner that it need be replaced with one of its redundant copies.
- **Information security:** To limit the potentially harmful effects of unauthorized access, the Internet Archive controls and monitors access privileges to the WARCs on its storage servers on a partner-by-partner basis. The abilities to access, manage, and download any collecting organization's WARC files is limited to Archive-It staff and the respective organization's authorized Archive-It software users. To further ensure their physical integrity, the files' data centers are located in access controlled, alarmed, and fire-protected buildings.

NYARC fully expects the above insurances to sufficiently protect and maintain uninterrupted access to its once ephemeral, web-native resources. However, it transcends and enhances them with one further layer of preservation and empowers its consortial institutions to do so as well. It automatically syncs all of its WARC files to the open-source and cloud-based digital repository management system DuraCloud.<sup>18</sup> While the Internet Archive maintains responsibility for the primary and secondary data centers that store and serve NYARC's and its other Archive-It partners' web archives for access, DuraCloud enables NYARC to manage separate *preservation* copies of its own WARC files, unmediated by any harvesting/access software service provider. These preservation stores are likewise redundant and geographically dispersed; management access is limited strictly to designated DuraCloud users authorized by NYARC; access, actions taken by authorized users, and each WARC's checksum value are routinely monitored should any event warrant the replacement of a damaged file. NYARC's member libraries also manage separate digital repository regimes specific to their respective institutional archiving mandates. Should any of these mandates expand to include the institution's or further web archives among

NYARC's collections, the web archiving programme manages their acquisition and regular update of WARC files from the original Internet Archive data stores.<sup>19</sup> In all respects, the NYARC web archiving programme's approach to storage and preservation adheres to the *LOCKSS* principle--"lots of copies keeps stuff safe."

### *Description and access*

From its inception, NYARC's investment in web archiving has been aligned with a commitment to provide wide and open access to all websites within the consortium's collections. Thus an important deliverable of NYARC's two-year Mellon-funded implementation grant was the development of sustainable and transferable metadata practices for describing the websites harvested and managed by the consortium with the Archive-It service. Staff at each NYARC library worked with metadata consultant Rebecca Guenther over several months to evaluate and test cataloging workflows in support of the development of guidelines for description.

Guenther's resulting documentation, 'Metadata application profile and a data dictionary for description of websites with archived versions,' was completed in June 2015. The profile details 'a rich record based on MARC' and the specific core elements recommended for the description of both live and archived versions of a website. The metadata profile contains information pertaining to the main fields and data dictionary of elements for description of sites with archived versions, as well as notes on MARC encoding and record samples. The metadata profile serves the purpose of offering specific guidance to staff at the NYARC libraries, but it is also 'cataloging rules agnostic,' and thus, extensible to the greater research community engaged in web archiving as a collection development activity.<sup>20</sup>

In keeping with the consortial web archiving workflow, NYARC's technical services staff begin their work in OCLC Connexion to create MARC records for live and archived websites. These records are made available to researchers in Arcade,<sup>21</sup> the collective online public access catalog (OPAC) of the three NYARC institutions, as well as in WorldCat. Another significant component of NYARC's recent grant was the procurement and implementation of a discovery layer solution. NYARC evaluated several discovery products available from commercial vendors and ultimately selected the Ex Libris Primo product,<sup>22</sup> with the objective of promoting greater

and more unified access to the consortium's Archive-It collections alongside NYARC's other rich holdings. NYARC's implementation, called NYARC Discovery, integrates Archive-It collection search results with many other material types in NYARC's collections, inclusive of books/e-books, journal articles, auction catalogs, newspaper articles, images, dissertations and photoarchives (see Fig. 3).<sup>23</sup> The consortium's Archive-It collection results are based on full-text indexing done by Archive-It at the time of each site's capture and routed to the NYARC Discovery interface via Archive-It's LOCKSS API.

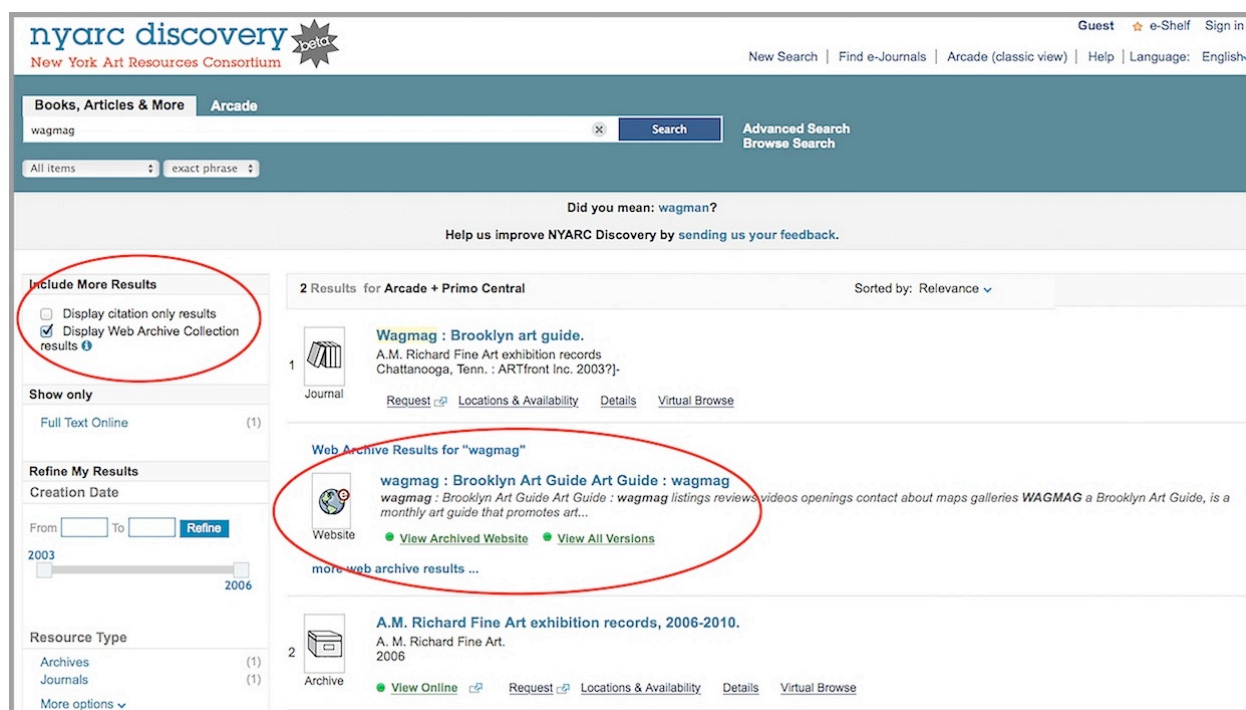


Fig. 3. A search in NYARC Discovery for the monthly Brooklyn Arts Guide 'Wagmag' yields web archive results from NYARC's Archive-It collections.

## Resources for collaboration among art librarians and web archivists

NYARC empowers and relies upon each of its member libraries to contribute in meaningful ways at one or more stages in the web archiving process in order to sustain its programme. It also, however, has benefitted and continues to benefit from resources available to consortia and independent programmes alike that wish to collaborate on web archiving's myriad technical, organizational, and intellectual challenges. As stewards of visual culture, art librarians are again

put in a unique position to both contribute to and benefit from the further development of these means to collect and preserve ephemera in such a richly diverse and expressive medium.

In its 2013 survey of web archiving programmes in the United States, the National Digital Stewardship Alliance (NDSA) reported that fully half of its respondents were actively participating or interested in participating in collaborative web archives.<sup>24</sup> NDSA is one, but not the only, network of libraries and archives in the United States that organizations can join or monitor in order to share web archiving ideas and resources. The Society of American Archivists' Web Archiving Roundtable connects more than 900 archivists and allied professionals through email lists, annual meetings, live and online educational resources, and more, in order to facilitate programme creation, resource sharing, and tool development/improvement.<sup>25</sup> A similarly missioned special interest group for art librarians involved or wishing to become involved in web archiving was proposed at the 2015 annual meeting of the Art Libraries Society of North America (ARLIS/NA) and will, at the authors' facilitation, meet for the first time at the 2016 joint annual meeting of ARLIS/NA and the Visual Resources Association (VRA). In the last year, NYARC's and other art librarians have taken the opportunities to meet at nationally scoped conferences and summits on web archiving unaffiliated to any specific professional organization, such as the *Web Archiving Collaboration: New Tools and Methods* meeting at Columbia University<sup>26</sup> and *Web Archives 2015: Capture, Curate, Analyze* meeting at the University of Michigan.<sup>27</sup> To stay connected in between such opportunities, NYARC has helped to form and to programme events for both New York metropolitan area Archive-It software service subscribers<sup>28</sup> and web archivists more broadly.<sup>29</sup>

## Conclusion

NYARC has been largely successful in establishing a rich and sustainable web archiving programme in the past two years due to its unique model, as a consortium eager to leverage the resources and expertise housed among its member institutions. It would not, however, have been able to achieve such strong growth in web archiving, much less in such a short period of time, had it not had access to an expansive collaborative network and had the innumerable formal and informal partnerships that have helped to develop the programme from its early stages (see Fig.

4). NYARC is grateful for the financial support of The Andrew W. Mellon Foundation and for the rich collaborative opportunity to serve as a host institution for NDSR New York. Partnership with the Pratt Institute School of Information graduate programme has been crucial to the success of NYARC's efforts to achieve quality web archive captures. The consortium is equally privileged to contribute to the strengthening of digital preservation knowledge and skillsets among Pratt's graduate students. Actively supportive and collaborative relationships with vendors and content producers have also been critical to the objectives of collecting, preserving, and providing access to born-digital art ephemera.

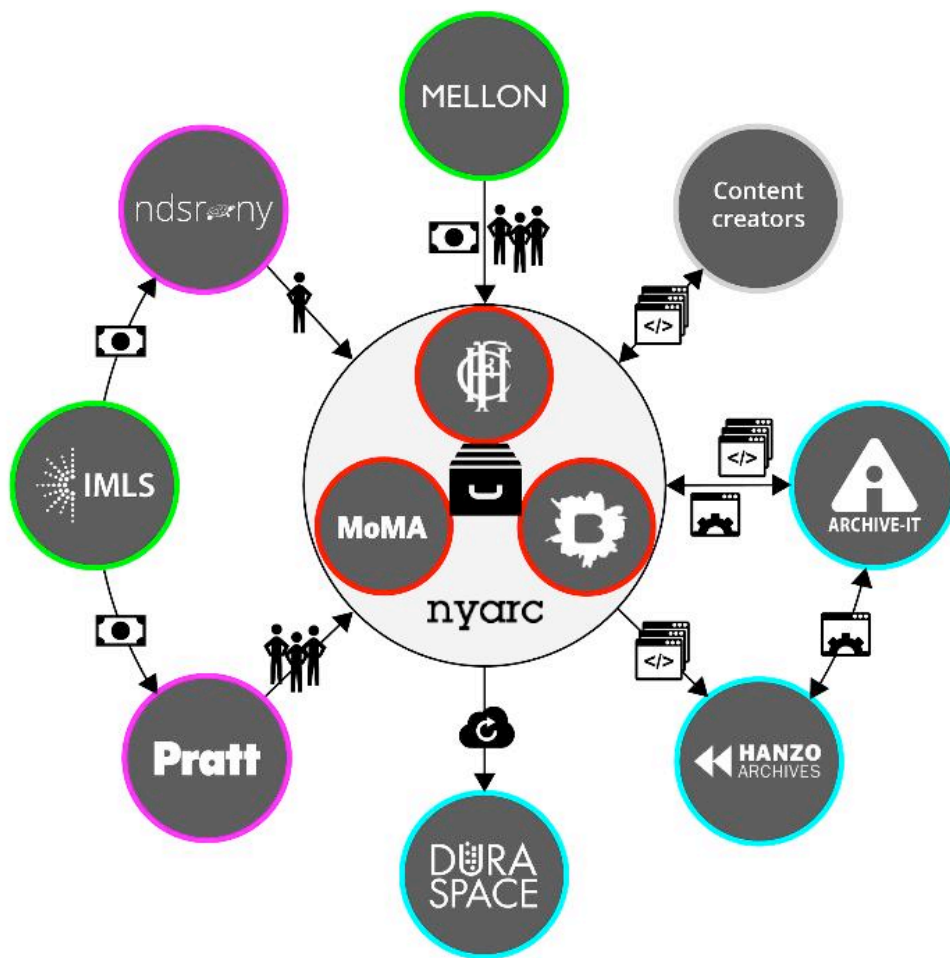


Fig. 4. Organizational diagram of institutions and their functions in NYARC's grant-funded 2013-15 web archiving programme.

It is evident that broader collaborations and partnerships will be necessary if we seek to truly harness the vast number of art-rich and ephemeral resources being produced on the web today and in the future. The NYARC libraries will continue to build upon their existing web archives, incrementally expanding collections and producing new access points to archived resources. Additionally, the consortium is in the early stages of expanding collaborative partnerships in web archiving to include institutions within the art and museum libraries community who possess complementary collecting objectives. NYARC is also invested in establishing new partnerships that will allow for greater understanding about the use of web archive collections through data and trend analysis, and in engaging our users to better grasp the unique challenges in the domain of art history.

## References

1. The Museum of Modern Art Library,  
[http://www.moma.org/learn/resources/library/faq\\_library\\_collection](http://www.moma.org/learn/resources/library/faq_library_collection)
2. Heather Slania, 'Online Art Ephemera: Web Archiving at the National Museum of Women in the Arts.' *Art Documentation: Journal of the Art Libraries Society of North America*. Vol. 32, No. 1 (Spring 2013). pp. 112-126.
3. NYARC, 'Three major New York City art libraries announce plans for a shared program to preserve digital art resources.'  
<http://www.nyarc.org/sites/default/files/NYARCborndigitalgrantOctober2013.pdf>
4. Archive-It, <https://archive-it.org>
5. NYARC Documentation: Web Archiving, <http://wiki.nyarc.org/web-archiving>
6. The National Digital Stewardship Residency: New York, <http://ndsr.nycdigital.org/>
7. NYARC Documentation: Web Archiving, 3. NYARC Collection Scope and Seed URL Nomination Process, <http://wiki.nyarc.org/web-archiving/3-nyarc-collection-scope-and-seed-url-nomination-process>
8. Columbia University Libraries Web Resources Collection Program,  
[https://library.columbia.edu/bts/web\\_resources\\_collection.html](https://library.columbia.edu/bts/web_resources_collection.html)
9. Wayback Machine, <http://archive.org/web/>

10. Corey Davis, 'Archiving the Web: A Case Study from the University of Victoria.' *code4lib Journal*, No. 26 (Fall 2014). Retrieved from <http://journal.code4lib.org/articles/10015>
11. Heather Slania, 'Online Art Ephemera: Web Archiving at the National Museum of Women in the Arts.' *Art Documentation: Journal of the Art Libraries Society of North America*. Vol. 32, No. 1 (Spring 2013). pp. 112-126
12. Hanzo Archives, <http://www.hanzoarchives.com/>
13. Karl-Rainer Blumenthal, 'Web Archive Management at NYARC: An NDSR Project Update.' *The Signal: Digital Preservation*. 2015. Retrieved from <http://blogs.loc.gov/digitalpreservation/2015/01/web-archive-management-at-nyarc-and-ndsr-project-update/>
14. NYARC Documentation: Web Archiving, 6. Quality Assurance (QA), <http://wiki.nyarc.org/web-archiving/quality-assurance>
15. ISO 28500:2009 Information and documentation — WARC file format, [http://www.iso.org/iso/iso\\_catalogue/catalogue\\_tc/catalogue\\_detail.htm?csnumber=44717](http://www.iso.org/iso/iso_catalogue/catalogue_tc/catalogue_detail.htm?csnumber=44717)
16. Archive-It Storage and Preservation Policy, <https://webarchive.jira.com/wiki/display/AITH/Archive-It+Storage+and+Preservation+Policy>
17. Paula De Stefano, Carl Fleischhauer, Andrea Goethals, Michael Kjörling, Nick Krabbenhoef, Chris Lacinak, Jane Mandelbaum, Kevin McCarthy, Kate Murray, Vivek Navale, Dave Rice, Robin Ruggaber, Trevor Owens, and Kate Zwaard, 'Checking Your Digital Content: What is Fixity, and When Should I be Checking It?' National Digital Stewardship Alliance. 2014. Retrieved from [http://digitalpreservation.gov/ndsa/working\\_groups/documents/NDSA-Fixity-Guidance-Report-final100214.pdf](http://digitalpreservation.gov/ndsa/working_groups/documents/NDSA-Fixity-Guidance-Report-final100214.pdf)
18. DuaCloud, <http://duracloud.org>
19. Karl-Rainer Blumenthal, 'Preserving NYARC's Web Archives: A Step Towards Long-term Stewardship.' 2015. Retrieved from <http://wiki.nyarc.org/web-archiving/long-term-storage-and-preservation>

20. Rebecca Guenther, 'Metadata Application Profile and Data Dictionary for Description of Websites with Archived Versions.' 2015. Retrieved from <http://wiki.nyarc.org/web-archiving/8-metadata-for-web-archives>
21. NYARC Arcade, <http://arcade.nyarc.org>
22. Ex Libris Primo, <http://www.exlibrisgroup.com/category/PrimoOverview>
23. NYARC Discovery, [http://nyarc-primo.hosted.exlibrisgroup.com/primo\\_library/libweb/action/search.do?vid=01NYARC](http://nyarc-primo.hosted.exlibrisgroup.com/primo_library/libweb/action/search.do?vid=01NYARC)
24. Jefferson Bailey, Abigail Grotke, Kristine Hanna, Cathy Hartman, Edward McCain, Christie Moffatt, & Nicholas Taylor, 'Web Archiving in the United States: A 2013 Survey.' National Digital Stewardship Alliance. 2014. Retrieved from [http://www.digitalpreservation.gov/ndsa/working\\_groups/documents/NDSA\\_USWebArchivingSurvey\\_2013.pdf](http://www.digitalpreservation.gov/ndsa/working_groups/documents/NDSA_USWebArchivingSurvey_2013.pdf)
25. Society of American Archivists Web Archiving Roundtable, <http://www2.archivists.org/groups/web-archiving-roundtable>
26. Web Archiving Collaboration: New Tools and Models, [https://library.columbia.edu/bts/web\\_resources\\_collection/Conferences/program.html](https://library.columbia.edu/bts/web_resources_collection/Conferences/program.html)
27. Web Archives 2015: Capture, Curate, Analyze, <http://www.lib.umich.edu/webarchivesconference>
28. Karl-Rainer Blumenthal, 'Web Archiving Happens Here: NYARC Hosts the First Meeting of Archive-It NY.' New York Art Resources Consortium. 2015. Retrieved from <http://www.nyarc.org/content/web-archiving-happens-here-nyarc-hosts-first-meeting-archive-it-ny>
29. Metropolitan New York Library Council (METRO) Web Archiving Special Interest Group, <http://libguides.metro.org/webarchiving>

Sumitra Duncan  
*NYARC Web Archiving Coordinator*  
*Frick Art Reference Library*  
*The Frick Collection*  
*10 East 71st Street*  
*New York, NY 10021*  
*USA*

*Email: [duncan@frick.org](mailto:duncan@frick.org)*

Karl-Rainer Blumenthal

*Web Archivist*

*Internet Archive*

*300 Funston Avenue*

*San Francisco, CA 94118*

*USA*

*Email: [karlb@archive.org](mailto:karlb@archive.org)*